PROSES KEPUTUSAN MARKOV

MAKALAH

Oleh

Drs. Yerizon, M. Si

NIP. 132051382

FAKULTAS PENDIDIKAN MATEMATIKA DAN IPA

IKIP PADANG

1998

BAB I
PENDAHULUAN

Dalam dunia usaha atau perusahaan, seorang pimpinan selalu dihadapkan pada situasi yang mengharuskan dia mengambil suatu tindakan dari beberapa kemungkinan tindakan yang ada. Misalnya dalam pemeliharaan sebuah mesin, jika ada bagian dari mesin tersebut bekerja kurang baik maka dia akan dihadapkan pada beberapa pilihan yaitu memperbaiki atau membiarkan saja atau mengganti dengan yang baru. Jika diperbaiki tentu akan diperlukan biaya dan produksi bisa meningkat. Sebaliknya jika tidak diperbaiki tentu tidak diperlukan biaya tetapi produksi berkurang. Dalam hal ini pimpinan tersebut harus menentukan suatu kebijaksanaan jangka panjang kapan dia memperbaiki, membiarkan atau mengganti bagian mesin tersebut sehingga produksi maksimal dan biaya rata-rata harian minimal.

Untuk menghitung biaya rata-rata harian yang minimal tentunya pimpinan tersebut akan menghitung biaya dari semua kemungkinan kebijaksanaan yang ada dan memilih biaya yang paling kecil. Jika dilihat contoh di atas maka keadaan mesin ada 3 yaitu baik, kurang baik, rusak dan kemungkinan tindakan yang akan diambil ada 2 yaitu memperbaiki atau tidak memperbaiki. Jadi pimpinan memiliki $2^3 = 8$ kemungkinan kebijaksanaan yang ada. Salah satu contoh dari 8 kebijaksanaan tersebut adalah jika mesin baik tidak diperbaiki , jika kurang baik diperbaiki dan jika rusak juga diperbaiki.

Selanjutnya jika kita punya N keadaan maka kita punya 2^N kemungkinan kebijaksanaan. Jika kita punya N keadaan dan 4 kemungkinan tindakan yang bisa diambil maka dia punya 4^N kemungkinan kebijaksanaan. Jumlah ini akan bertambah dengan cepat jika ada lebih dari 4 tindakan yang mungkin. Dengan jumlah ini tentunya dia akan kesulitan untuk menghitung semua biaya rata-rata harian dari kebijaksanaan tersebut dan menentukan biaya rata-rata harian minimalnya.

Untuk memudahkan menghitung biaya rata-rata harian minimal dari masalah di atas bisa digunakan Proses Keputusan Markov. Dalam proses ini keputusan yang diambil sekarang hanya bergantung pada keputusan yang terakhir diambil. Misalnya jika kita sekarang mengambil keputusan kelima maka kita hanya perlu memperhatikan keputusan keempat dan tidak memperhatikan keputusan ketiga, kedua dan pertama. Salah satu metode yang ada dalam Proses Keputusan adalah metode Iterasi-Policy (Iterasi-Kebijaksanaan) yang akan dibahas dalam makalah ini. Metode ini akan mengurangi banyaknya perhitungan, sehingga hasil yang diinginkan diperoleh sebelum biaya rata-rata harian semua kebijaksanaan dihitung.

BAB II

PERMASALAHAN

Berdasarkan uraian pada Pendahuluan maka yang menjadi masalah adalah sulitnya menghitung biaya rata-rata harian minimal dengan cara menghitung biaya rata-rata harian untuk semua kebijaksanaan yang mungkin jika jumlah kebijaksanaan yang mungkin terlalu banyak.

Selanjutnya yang menjadi pertanyaan adalah apakah ada cara atau metode yang efektif dan efisien untuk menghitung biaya rata-rata harian minimal dengan tanpa menghitung semua kemungkinan yang ada?.

BAB III

PEMBAHASAN

A. TEORI PENDUKUNG

1. Rantai Markov Parameter Diskrit

Definisi 1.1

Misal $\{X_n, n=0,1,2,3,...\}$ adalah barisan variabel acak dengan ruang keadaan diskrit I.

Untuk n = 0,1,2,...

X_n = keadaan sistem pada waktu n

Proses stokastik $\{X_n, n=0,1,2,...\}$ disebut *Rantai Markov Parameter Diskrit* jika untuk tiap n = 0,1,2,...

$$P\{X_{n+1}=i_{n+1} | X_0=i_0,...,X_n=i_n\} = P\{X_{n+1}=i_{n+1} | X_n=i_n\} \qquad (1)$$

untuk setiap $i_0,...,i_{n+1}$ (Ross, 1983 :100).

Untuk selanjutnya persamaan (1) disebut *Sifat Markov*.

Definisi 1.2

Misal $\{X_n, n=0,1,2,...\}$ adalah rantai Markov dengan waktu diskrit. Untuk setiap $n \geq m \geq 0$ dan keadaan I dan j didefinisikan;

a. Fungsi massa peluang

$$P_j(n) = P[X_n = j] \qquad (2)$$

b. Fungsi massa peluang bersyarat

$$P_{ij}(m,n) = P[X_n = j | X_m = I] \qquad (3)$$

Persamaan (3) dinamakan *fungsi peluang transisi* dari rantai Markov (Parzen, 1962 : 193).

Definisi 1.3.

Rantai Markov dinamakan *homogen* jika $p_{jk}(m,n)$ hanya bergantung pada n-m dan tidak bergantung pada m dan n. Untuk selanjutnya kita hanya akan membahas rantai Markov homogen, dengan

$$P_{ij}(n) = P\{X_{n+t} = j | X_t = i\} \qquad (4)$$

Persamaan (4) disebut *peluang transisi n-langkah* (Ross, 1983 :103).

Sifat: - $p_{ij} \geq 0$ untuk $i,j \geq 0$

- $\sum_{j=0}^{\infty} p_{ij} = 1$, $i = 0, 1, 2, ...$

2. Persamaan Chapman-Kolmogorov

Misal $\{X_n, n=0,1,2,...\}$ rantai Markov homogen dengan ruang keadaan I. Untuk setiap n, peluang transisi $p_{ij}(n)$ memenuhi

$$p_{ij}(n+m) = \sum_{k=0}^{\infty} p_{ik}(n)p_{kj}(m) \quad \forall i,j \in I \text{ dan } \forall m,n \geq 0 \qquad (5)$$

(Parzen, 1962 : 194)

Bukti.

$$p_{ij}(n+m) = P\{X_{n+m} = j \mid X_0 = i\}$$

$$= \sum_{k=0}^{\infty} P\{X_{n+m} = j, X_n = k \mid X_0 = I\}$$

$$= \sum_{k=0}^{\infty} P\{X_{n+m} = j \mid X_n = k, X_0 = i\} P\{X_n = k \mid X_0 = i\}$$

$$= \sum_{k=0}^{\infty} p_{kj}(m)p_{ik}(n)$$

Untuk m+n = 0, didefinisikan

$$p_{ij}(0) = 1, \text{ jika } i = j$$
$$= 0, \text{ jika } i \neq j$$

Catatan.

Jika $P^{(1)} = P = [\, p_{ij}\,]$ maka $P^{(n)} = P^{(1)}P^{(n-1)} = P^{(2)}P^{(n-2)} = \ldots = P^{(n)}$

3. Klasifikasi Keadaan

Definisi 3.1

Misal $\{X_n, n=0,1,2,...\}$ adalah rantai Markov. Keadaan j dikatakan dapat dicapai dari keadaan i jika terdapat $n \in N$ sehingga $p_{ij}(n) > 0$ dan dinotasikan dengan $i \to j$. Jika $i \to j$ dan $j \to i$, maka keadaan i dan j dikatakan *berkomunikasi* dan dinotasikan dengan $i \leftrightarrow j$ (Ross, 1983 :104).

Catatan.

- Relasi komunikasi adalah relasi ekivalen.
- Himpunan keadaan yang berkomunikasi dengan j ditulis

$$C(j) = \{ k \in I \mid j \leftrightarrow k \}.$$

- Suatu $C \subset I$ disebut *klas komunikasi* jika ada $j \in I$ sehingga $C \subset C(j)$.

Definisi 3.2

Sebuah rantai Markov dikatakan *tak terreduksi* jika hanya terdapat satu klas komunikasi (Ross, 1983 : 104).

Definisi 3.3

Keadaan i dikatakan *mempunyai periode d(i)* jika d(i) adalah faktor persekutuan terbesar dari $n \in N$ sehingga $p_{ii}(n) > 0$. Jika $d(i) = 1$ maka keadaan i adalah *aperiodik* dan jika $d(i) > 1$ maka keadaan i adalah *periodik*.

Kita definisikan *peluang pertama kali sampai dari keadaan i ke keadaan j* dengan

$$f_{ij}(n) = P\{X_n = j, X_r \neq j, 1 \leq r \leq n-1 \mid X_0 = i\}$$

untuk sebarang $i,j \in I$ (Osaki, 1992 :46).

Catatan.

- $f_{ij}(0) = 0$
- $f_{ij}(1) = p_{ij}$

Selanjutnya kita definisikan *peluang transisi terakhir dari keadaan i ke keadaan j* dengan f_{ij} dimana $f_{ij} = \sum_{n=1}^{\infty} f_{ij}(n)$

Definisi 3.4

Jika $f_{ii} = 1$ maka keadaan i dikatakan *recurrent.*

Jika $f_{ii} < 1$ maka keadaan i dikatakan *transient.*

Misal μ_{jj} adalah ekspektasi banyak transisi dibutuhkan untuk kembali ke keadaan j, yaitu

$$\mu_{jj} = \infty, \text{jika j transient}$$

$$= \sum_{n=1}^{\infty} nf_{jj}, \text{jika j recurrent}$$

(Osaki, 1992 : 51).

4. Sifat Jangka Panjang dari Rantai Markov

Definisi 4.1

Sebuah rantai Markov $\{X_n, n=0,1,2,...\}$ dengan ruang keadaan I dikatakan *mempunyai distribusi jangka panjang* jika terdapat distribusi peluang $\{ \pi_j, j \in I \}$ yang mempunyai sifat untuk semua i dan j di I berlaku,

$$p_{ij}(n) \rightarrow \pi_j \text{ jika } n \rightarrow \infty$$

(Parzen, 1962 :219).

Teorema 4.1

Untuk setiap rantai Markov tak terreduksi dengan ruang keadaan I, terdapat barisan $\{ \pi_j, j \in I \}$ sehingga untuk setiap i dan j di I berlaku

$$\frac{1}{n}\sum_{k=1}^{n} p_{ij}(k) \rightarrow \pi_j > 0 \text{ jika } n \rightarrow \infty. \qquad (6)$$

Persamaan (6) menyatakan bahwa barisan $p_{ij}(n)$ konvergen dalam rata-rata Cesaro ke sebuah limit π_j yang bebas dari i. Jika I hingga maka $\{ \pi_j, j \in I \}$ adalah distribusi peluang, yaitu $\sum_{j \in I} \pi_j = 1$

Teorema 4.2

Jika sebuah rantai Markov positif recurrent dan aperiodik maka terdapat limit peluang

$$p_{ij}(n) \rightarrow \pi_j \text{ jika } n \rightarrow \infty. \ (i,j = 0,1,2,...)$$

yang bebas dari keadaan awal i , dimana $\{ \pi_j, j \in I \}$ adalah jawab tunggal dan positif untuk sistem persamaan

$$\pi_j = \sum_{i=0}^{\infty} \pi_i p_{ij} \qquad (j = 0,1,2,...)$$

$$\sum_{j=0}^{\infty} \pi_j = 1$$

dan disebut *distribusi stasioner* untuk sebuah rantai Markov (Chung, 1967 : 35).

Ekspektasi Biaya Rata-rata Per-satuan Waktu

Asumsi.

Ekspektasi biaya total pada waktu pertama melewati masa 0 sehingga proses mencapai keadaan transisi r adalah hingga untuk setiap keadaan awal $X_0 = i$.

Misalkan biaya $C(X_t)$ diadakan bila proses dalam keadaan X_t pada waktu t untuk t = 0, 1, 2, ... $C(X_t)$ adalah variabel random yang mempunyai nilai salah satu dari C(0), C(1), ..., C(M) dan fungsi C(.) bebas dari t. Maka ekspektasi biaya rata-rata yang diadakan pada n periode pertama diberikan dengan ekspresi

$$E\left[\frac{1}{n} \sum_{t=1}^{n} C(X_t) \right]$$

Dengan menggunakan teorema 4.1 maka diperoleh *biaya rata-rata jangka panjang per-satuan waktu* yaitu

$$\lim_{n\to\infty} \left\{ E\left[\frac{1}{n} \sum_{t=1}^{n} C(X_t) \right] \right\} = \sum_{j\in I} \pi_j C(j)$$

B. MODEL KEPUTUSAN MARKOV PARAMETER DISKRIT

Misalkan sebuah sistem dinamik (yaitu suatu sistem yang dapat dijadikan persoalan program dinamis) diperiksa pada waktu n = 0, 1, 2, ... dengan jarak antar pemeriksaan sama. Pada setiap pemeriksaan sistem tersebut diklasifikasikan kedalam suatu keadaan dari sejumlah keadaan yang mungkin, kemudian keputusan atau tindakan diambil untuk memperbaiki sistem tersebut.

Misalkan I adalah himpunan keadaan yang mungkin. Untuk setiap $i \in I$, misalkan A(i) adalah himpunan aksi yang mungkin.Dalam hal ini I dan A(i) diambil berhingga. Sebagai konsekuensi dari pengambilan keputusan tersebut adalah diperlukan biaya.

Sistem dinamik yang dikontrol ini disebut *Model Keputusan Markov Parameter Diskrit* bila sifat Markov dipenuhi (Tijms, 1986 : 161).

Jika pada waktu pengambilan keputusan, aksi $a \in A(i)$ dipilih pada keadaan i, maka yang berikut akan terjadi

(1) Biaya $C_i(a)$ diadakan

(2) Pada pengambilan keputusan berikutnya sistem akan berada dalam keadaan j dengan peluang $p_{ij}(a)$ dimana

$$\sum_{j \in I} p_{ij}(a) = 1 \, , \ i \in I$$

Dalam masalah ini $C_i(a)$ menyatakan ekspektasi biaya yang diadakan sebelum masa keputusan berikutnya, bila aksi a dipilih pada keadaan i. Perlu ditegaskan bahwa pemilihan ruang keadaan dan himpunan aksi biasanya tergantung pada struktur biaya dari masalah tersebut.

Selanjutnya akan diperkenalkan beberapa konsep yang akan dibutuhkan dalam algoritma policy-iterasi.

Sebuah aturan atau policy dalam pengontrolan suatu sistem adalah suatu petunjuk untuk pengambilan aksi pada setiap waktu pemeriksaan. Pada dasarnya sebuah aturan kontrol mungkin sangat komplit dalam arti bahwa penentuan aksi mungkin bergantung pada seluruh sistem. Tapi dengan sifat Markov di atas dan waktu perencanaan takhingga panjangnya, maka kita hanya akan mempertimbangkan apa yang disebut *policy* *stasionary*. Sebuah policy stasionary R adalah himpunan peraturan yang tertentu dan selalu menentukan aksi tunggal R_i bila sistem dalam keadaan i pada waktu pengambilan keputusan. Kita punya hubungan $R_i \subseteq R$ dan $a \in R_i$.

Definisikan untuk n = 0, 1, 2, ...

$$X_n = \text{keadaan dari sistem pada waktu keputusan ke-n}$$

Dibawah policy stasionary R yang diberikan, kita punya

$$P\{X_{n+1} = j \mid X_n = i\} = p_{ij}(R_i)$$

tanpa memperhitungkan sistem sampai waktu n. Jadi dibawah policy stasionary R proses stokastik $\{X_n, n = 0,1,2,...\}$ adalah rantai Markov dengan waktu diskrit, dengan peluang transisi satu langkah $p_{ij}(R_i)$.Juga rantai Markov ini menyediakan biaya $C_i(R_i)$ untuk tiap waktu sistem mengunjungi keadaan i.

Selanjutnya peluang transisi n-langkah dari rantai Markov $\{X_n, n = 0,1,2,...\}$ adalah

$$p_{ij}^{(n)}(R) = P\{X_n = j \mid X_0 = i\}, \; i,j \in I \text{ dan } n = 1, 2, 3, ...$$

dimana $p_{ij}^{(1)}(R) = p_{ij}(R_i)$ dan juga memenuhi

$$p_{ij}^{(n+1)}(R) = \sum_{k \in I} p_{ik}^{(n)}(R)\, p_{kj}(R_k)$$

(Tijms, 1986 :164).

Dalam aplikasi praktis setiap policy stasionary biasanya punya sifat *unichain* yaitu terdapat suatu keadaan yang dapat dicapai dari sebarang keadaan dibawah policy tersebut. Untuk itu diperlukan asumsi berikut.

Asumsi 3.1

Untuk setiap policy stasionary R, terdapat sebuah keadaan r (mungkin bergantung pada R) yang dapat dicapai dari sebarang keadaan dibawah policy R.

Dibawah asumsi 3.1 maka untuk setiap policy stasionary R kita punya

$$\pi_j(R) = \lim_{m \to \infty} \frac{1}{m} \sum_{n=1}^{m} p_{ij}^{(n)}(R)$$

ada dan bebas dari keadaan awal $X_0 = i$.

Distribusi stasionary $\{\ \pi_j, j \in I\ \}$ memenuhi sistem persamaan linear berikut

$$\pi_j(R) = \sum_{k \in I} p_{kj}(R_k)\, \pi_k(R)$$

$$\sum_{j \in I} \pi_j(R) = 1 \qquad (7)$$

Sistem persamaan ini punya jawab tunggal. Selanjutnya dengan relasi (7) kita punya dengan peluang 1, *biaya rata-rata jangka panjang persatuan waktu* bila menggunakan policy R ,katakan g(R) dengan

$$g(R) = \sum_{j \in I} C_j(R_j)\, \pi_j(R) \qquad (8)$$

Definisi 3.1

Sebuah policy stasionary R* dikatakan *policy biaya rata-rata optimal* jika $g(R^*) \leq g(R)$ untuk setiap policy stasionary R (Tijms, 1986 : 165).

Dengan perhitungan biasa sangat sulit untuk menemukan policy optimal biaya rata-rata dengan menghitung biaya rata-rata untuk setiap policy stasionary dari (7) dan (8). Misalnya dalam kasus ruang keadaan mempunyai N keadaan dan untuk setiap aksi memuat 2 aksi maka jumlah kemungkinan policy stasionary adalah 2^N. Bila jumlah aksi 4 maka jumlah policy stasionary yang mungkin menjadi 4^N. Walaupun demikian ada sebuah algoritma yang efisien untuk menemukan policy stasionary optimal yang akan diberikan. Algoritma ini mengkonstruksi barisan policy yang diperbaiki sebelum policy optimal biaya rata-rata ditemukan. Dalam memperbaiki sebuah policy R yang diberikan, diperlukan apa yang disebut *nilai relatif* dari bermacam keadaan awal bila policy R digunakan. Nilai relatif tersebut menunjukkan akibat sementara dari keadaan awal pada total ekspektasi biaya dibawah policy yang diberikan.

1. Nilai Relatif Berkaitan dengan Sebuah Policy R yang Diberikan

Ekspektasi biaya yang diadakan pada waktu keputusan ke-t dengan $X_0 = i$ dan policy R digunakan adalah

$$\sum_{j \in I} p_{ij}^{(t)}(R)\, C_j(R_j)$$

Maka *ekspektasi biaya total* pada n waktu keputusan pertama bila $X_0 = i$ dan policy R digunakan adalah

$$V_n(i,R) = \sum_{t=0}^{n-1} \sum_{j\in I} p_{ij}^{(t)}(R)\, C_j(R_j)$$

Selanjutnya kita peroleh,

$$\lim_{n\to\infty} \frac{Vn(i,R)}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} \sum_{j\in I} p_{ij}^{(t)}(R)\, C_j(R_j)$$

$$= \sum_{j\in I} C_j(R_j) \lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} p_{ij}^{(t)}(R)$$

$$= \sum_{j\in I} C_j(R_j)\, \pi_j(R)$$

$$= g(R)$$

Dari relasi ini, maka terdapat nilai bias $v_i(R)$, sehingga untuk setiap

$$i \in I,\ Vn(i,R) \approx n\, g(R) + v_i(R) \text{ untuk n besar} \qquad (9)$$

Catatan.

$v_i(R) - v_j(R) \approx Vn(i,R) - Vn(j,R)$ dimana

$v_i(R) - v_j(R)$ adalah selisih ekspektasi biaya total bila mulai dalam keadaan I dan selanjutnya dalam keadaan j dengan menggunakan aturan R.

Sebelum kita membuktikan bahwa biaya rata-rata g(R) dan nilai relatif $v_i(R)$, $i \in I$ memenuhi sebuah sistem persamaan linear, kita perhatikan dulu persamaan berikut

$$Vn(i,R) = C_i(R_i) + \sum_{j\in I} p_{ij}(R_i)\, Vn\text{-}1(j,R) \quad n \geq 1 \text{ dan } i \in I$$

Dengan mensubtitusikan (9) ke dalam persamaan ini diperoleh

$$g(R) + v_i(R) \approx C_i(R_i) + \sum_{j\in I} p_{ij}(R_i)\, v_j(R)$$

Misal r suatu keadaan tertentu yang dapat dicapai dari sebarang keadaan dengan menggunakan policy R, setelah berhingga banyak transisi.

Selanjutnya definisikan untuk tiap keadaan $i \in I$

Tir (R) = ekspektasi waktu sampai pertama kali ke keadaan r bila dimulai pada keadaan i dan menggunakan aturan R.

Kir(R) = ekspektasi biaya yang diadakan sampai pertama kali ke keadaan r bila dimulai dalam keadaan i dan menggunakan aturan R.

Berikutnya kita definisikan

$$w_i(R) = Ki(R) - Kr(R)\, Ti(R) / Tr(R) \tag{10}$$

sebagai kosekwensi dari (10) diperoleh

$$w_r(R) = 0$$

Dalam teorema berikut kita akan membuktikan bahwa biaya rata-rata persatuan waktu dan nilai relatif dapat dihitung secara simultan dengan penyelesaian sebuah sistem persamaan linear.

Teorema 3.1

Misalkan policy stasionary R diberikan

Misal g dan v_i adalah sebarang solusi dari,

$$x_i = C_i(R_i) - y + \sum_{j\in I} p_{ij}(R_i)\, x_j \tag{11}$$

(a). Biaya rata-rata g(R) dan nilai relatif w_i (R), $i \in I$ memenuhi (11) dan $g = g(R)$ serta untuk suatu konstanta k, $v_i = w_i(R) + k$ untuk setiap $i \in I$.

(b). Misal s sebarang keadaan dengan $v_s = 0$. Maka (11) bersama $v_s = 0$ mempunyai jawab tunggal (Tijms, 1986 : 167).

Bukti.

(a). Dari definisi Tir(R) dan Kir(R) kita dapat tuliskan,

$$Tir(R) = 1 + \sum_{j\in I, j\neq r} p_{ij}(R_i)\; Tjr(R)$$

$$Kir(R) = C_i(R_i) + \sum_{j\in I, j\neq r} p_{ij}(R_i)\, Kjr(R)$$

Maka,

$$Kir(R) - g(R)\, Tir(R) = C_i(R_i) + \sum_{j\in I, j\neq r} p_{ij}(R_i)\, Kjr(R) -$$

$$g(R)\,(1 + \sum_{j\in I, j\neq r} p_{ij}(R_i)\; Tjr(R))$$

$$= C_i(R_i) - g(R) + \sum_{j\in I, j\neq r} p_{ij}(R_i)\{Kjr(R) - g(R)\,Tjr(R)\}$$

Karena $w_i(R) = Kir(R) - g(R)\,Tir(R)$ maka

$$w_i(R) = C_i(R_i) - g(R) + \sum_{j\in I} p_{ij}(R_i)\, w_i(R),$$

Selanjutnya,

$$\text{Klaim: } v_i = \sum_{t=0}^{n-1} \sum_{j\in I} p_{ij}^{(t)}(R) C_j(R_j) - mg + \sum_{j\in I} p_{ij}^{(m)}(R) v_j \text{ untuk } m = 1, 2, 3, \ldots \quad (12)$$

Bukti Klaim.

Akan dibuktikan dengan induksi atas m.

Untuk m = 1, maka

$$v_i = \sum_{t=0}^{0} \sum_{j\in I} p_{ij}^{(t)}(R)\, C_j(R_j) - g + \sum_{j\in I} p_{ij}^{(1)}(R_i)\, v_j$$

$$= \sum_{j\in I} p_{ij}^{(0)}(R)\, C_j(R_j) - g + \sum_{j\in I} p_{ij}(R_i)\, v_j$$

$$= C_i(R_i) - g + \sum_{j\in I} p_{ij}(R_i)\, v_j$$

Catatan.

$p_{ij}^{(0)}(R) = 1$, jika $i = j$

$= 0$, jika $i \neq j$

Jadi benar untuk m = 1.

Andaikan benar untuk m = n.

Selanjutnya untuk m = n+1, perhatikan yang berikut.

Substitusikan (11) ke sisi kanan dari (12) maka diperoleh

$$v_i = \sum_{t=0}^{n-1} \sum_{j\in I} p_{ij}^{(t)}(R)\, C_j(R_j) - n\,g + \sum_{j\in I} p_{ij}^{(n)}(R)\ \{C_j(R_j) - g + \sum_{j\in I} p_{jk}(R_j)\, v_k\}$$

$$= \sum_{t=0}^{n-1} \sum_{j\in I} p_{ij}^{(t)}(R)\, C_j(R_j) + \sum_{j\in I} p_{ij}^{(n)}(R)\ C_j(R_j) - n\,g - g\sum_{j\in I} p_{ij}^{(n)}(R) +$$

$$\sum_{j\in I} p_{ij}^{(n)}(R) \sum_{j\in I} p_{jk}(R_j)\, v_k$$

$$= \sum_{t=0}^{n} \sum_{j\in I} p_{ij}^{(t)}(R)\, C_j(R_j) - (n+1)\, g + \sum_{k\in I} \{ \sum_{j\in I} p_{ij}^{(n)}(R)\, p_{jk}(R_j) \}\, v_k$$

$$= \sum_{t=0}^{n} \sum_{j\in I} p_{ij}^{(t)}(R)\, C_j(R_j) - (n+1)\, g + \sum_{k\in I} p_{ik}^{(n+1)}(R)\, v_k$$

Karena benar untuk $m = n+1$, maka (12) benar untuk setiap $m = 1,2,3,..$

Selanjutnya (11) dapat ditulis

$$v_i = Vm(i,R) - m\, g + \sum_{j\in I} p_{ij}^{(m)}(R)\, v_j\,, \quad i \in I \tag{13}$$

Dengan membagi kedua ruas dari (13) dengan m dan mengambil limitnya untuk $m \to \infty$ diperoleh $g = g(R)$. Untuk membuktikan bagian kedua dari pernyataan (a) kita misalkan $\{g, v_i\}$ dan $\{ g, v_i' \}$ adalah dua solusi dari (12), maka dari (13) diperoleh

$$v_i - v_i' = \sum_{j\in I} p_{ij}^{(m)}(R)\, \{v_j - v_j'\} \qquad i \in I\,, m \geq 1 \tag{14}$$

Dengan menjumlahkan kedua sisi (14) atas $m = 1,2,\ldots, n$ dan selanjutnya dibagi dengan n diperoleh,

$$v_i - v_i' = \sum_{j\in I} \{\frac{1}{n} \sum_{m=1}^{n} p_{ij}^{(m)}(R)\}\{v_j - v_j'\} \qquad i \in I\,, n \geq 1$$

Dengan mengambil limitnya untuk $n \to \infty$ dan menggunakan (3.2) diperoleh

$$v_i - v_i' = \sum_{j\in I} \pi_j(R)\, \{v_j - v_j'\} \qquad i \in I$$

Karena sisi kanan tidak bergantung pada I, maka pernyataan benar dengan mengambil $v_i' = w_i(R)$ dan $k = \sum_{j\in I} \pi_j(R)\, \{v_j - v_j'\}$.

(b). Misalkan $\{ g, v_i \}$ sebarang solusi dari (11). Karena $\sum_{j\in I} p_{ij}(R_i) = 1$ untuk setiap

$i \in I$, maka untuk sebarang konstanta γ bilangan g dan $v_i' = v_i + \gamma$ memenuhi (11). Karena itu persamaan (11) bersama dengan $v_s = 0$ untuk suatu keadaan s mempunyai solusi. Mengingat pernyataan (a) maka solusi ini tunggal.

2. Algoritma Iterasi-Policy

Teorema berikut sangat penting dalam Algoritma Iterasi-Policy.

Teorema 3.2

Misalkan g dan v_i, $i \in I$ bilangan yang diberikan. Misalkan policy $\underline{R}$ mempunyai sifat

$$C_i(\underline{R}_i) - g + \sum_{j \in I} p_{ij}(\underline{R}_i)\, v_j \le v_i, \text{ untuk setiap } i \in I \tag{15}$$

Maka $g(\underline{R}) \le g$ (16)

Teorema ini juga benar bila tanda pada (15) dan (16) dibalik (Tijms, 1986 : 170).

Bukti.

Kita dapat membuktikan dengan induksi atas m seperti bukti (12) bahwa,

$$Vm(i,\underline{R}) - m\, g + \sum_{j \in I} p_{ij}^{(m)}(\underline{R})\, v_j \le v_i \text{ untuk } i \subset I, m = 1, 2, 3, \ldots \tag{17}$$

Selanjutnya bagi kedua ruas (17) dengan m dan untuk $m \to \infty$ diperoleh $g(\underline{R}) \le g$. Jelas dari bukti di atas bahwa teorema ini juga benar jika tanda dari (15) dan (16) dibalik.

Algoritma Iterasi-Policy

Langkah 0.

Pilih sebuah policy stasionary R.

Langkah 1 (Langkah penentuan nilai).

Untuk sebarang policy R hitung solusi tunggal $\{ g(R), v_i(R) \}$ dari sistem persamaan linear

$$x_i = C_i(R_i) - y + \sum_{j \in I} p_{ij}(R_i)\, x_j$$

$$x_s = 0$$

dimana s keadaan sebarang yang dipilih.

Langkah 2 (Langkah memperbaiki policy)

Untuk tiap keadaan $i \in I$, tentukan sebuah aksi a yang memenuhi minimum dalam

$$\min_{a \in A(i)} \{ C_i(\{a\}) - g(R) + \sum_{j \in I} p_{ij}(\{a\})\, v_j(R) \} \tag{18}$$

Policy stasionary baru $\underline{R}$ diperoleh dengan memilih $\underline{R}_i = \{a\}$ yang diperoleh dari (18) dengan kesepakatan bahwa $\underline{R}_i$ dipilih sama dengan aksi yang lama R_i bila aksi ini menghasilkan minimum di (18).

Langkah 3 (Uji kekonvergenan).

Jika policy baru $\underline{R}$ sama dengan policy yang lama R, maka algoritma dihentikan R. Untuk yang lainnya kembali ke langkah 1 dengan R diganti $\underline{R}$ (Tijms, 1986 :171)

Teorema 3.3.

Jika algoritma di atas konvergen ke suatu policy stasionary R* maka R* adalah biaya rata-rata optimal (Tijms, 1986 :172).

Bukti.

Dari langkah 2 kita punya untuk setiap $i \in I$,

$$\min_{a \in A(i)} \{ C_i(\{a\}) - g(R^*) + \sum_{j \in I} p_{ij}(\{a\})\, v_j(R^*) \}$$

$$= C_i(R_i^*) - g(R^*) + \sum_{j \in I} p_{ij}(R_i^*)\, v_j(R^*) = v_i(R^*) \tag{19}$$

Akibat (19) maka untuk sebarang policy R yang dipilih,

$$C_i(R_i) - g(R^*) + \sum_{j \in I} p_{ij}(R_i)\, v_j(R^*) > v_j(R^*) \text{ untuk setiap } i \in I.$$

Maka menurut teorema 3.2 diperoleh $g(R) \geq g(R^*)$ untuk setiap R. Ini menunjukkan bahwa policy R adalah biaya rata-rata optimal.

Contoh.

Misalkan sebuah peralatan diperiksa setiap hari untuk mengetahui kondisi terakhirnya. Misalkan kondisi kerjanya i = 1, 2, 3, 4, 5. Dalam hal ini kondisi kerja i lebih baik dari i+1. Jika pada keadaan i sistem tidak diperbaiki, maka pada hari berikutnya keadaannya menjadi j dengan peluang q_{ij}. Diasumsikan peralatan tersebut tidak bisa memperbaiki dirinya sendiri, sehingga $q_{ij} = 0$ jika $j < i$. Pada keadaan i = 5 peralatan tersebut tidak bisa lagi dipakai dan membutuhkan dua hari untuk perbaikan dengan biaya 1 perhari. Setelah itu keadaannya menjadi i = 1. Karena kita memeriksa setiap hari dan waktu perbaikan dua hari, maka kita butuh satu keadaan tambahan. Maka himpunan keadaan menjadi I = { 1,2,3,4,5,6 }. Kita akan menentukan suatu kebijakan sehingga biaya rata-rata jangka panjangnya minimum. Misal q_{ij} adalah peluang menurunnya kondisi kerja peralatan tersebut dengan data pada tabel 3.1.

Tabel 3.1. Peluang Penurunan Kondisi Kerja (q_{ij})

i \ j	1	2	3	4	5
1	0,75	0,20	0,05	0,00	0,00
2	0,00	0,50	0,20	0,20	0,10
3	0,00	0,00	0,50	0,25	0,25
4	0,00	0,00	0,00	0,30	0,70

Dari persoalan di atas kita peroleh,

A(i) = { 0, 1 } dimana a = 1, jika peralatan diperbaiki

= 0, lainnya

$C_i(1) = 1, C_i(0) = 0.$

$A(1) = \{0\}, A(i) = \{0,1\}$ untuk $1 < i < 5$

$A(5) = A(6) = \{1\}$

$P_{i1}(\{1\}) = 1$ untuk $1 < i < 5,$

$P_{56}(\{1\}) = p_{61}(\{1\}) = 1$

$P_{ij}(\{0\}) = q_{ij}$, untuk $1 \le i < 5, \quad i \le j$

$P_{ij}(\{a\}) = 0$, selain dari ketentuan di atas

Akan dicari policy stasionary R* yang menghasilkan biaya rata-rata optimal dengan menggunakan algoritma iterasi-policy.

Iterasi 1.

Langkah 1.

Misal $R^{(1)}$ policy terakhir dengan aksi

$R_1^{(1)} = 0, R_2^{(1)} = R_3^{(1)} = R_4^{(1)} = R_5^{(1)} = R_6^{(1)} = 1$

Biaya rata-rata $g(R^{(1)})$ dan nilai relatif $v_i(R^{(1)})$ dapat dihitung sebagai solusi tunggal dari sistem persamaan linear

$v_1 = 0 - g + 0{,}75\, v_1 + 0{,}20\, v_2 + 0{,}05\, v_3$

$v_k = 1 - g + \quad v_1 \;, k = 2, 3, 4$

$v_5 = 1 - g + v_6$

$v_6 = 1 - g + v_1$

$v_6 = 0$

dimana s = 6 diambil untuk $v_s = 0$. Solusi dari sistem persamaan linear tersebut adalah

$g(R^{(1)}) = 0{,}2 \quad v_1(R^{(1)}) = -0{,}8 \quad v_2(R^{(1)}) = 0 \quad v_3(R^{(1)}) = 0$

$v_4(R^{(1)}) = 0 \quad v_5(R^{(1)}) = 0{,}8 \quad v_6(R^{(1)}) = 0$

Langkah 2.

Perhitungan untuk langkah memperbaiki policy bagi $R^{(1)}$ dapat dilihat pada tabel 3.2. Dari tabel 3.2 kita peroleh policy baru $R^{(2)}$ yang punya aksi

$R_1^{(2)} = R_2^{(2)} = 0, R_3^{(2)} = R_4^{(2)} = R_5^{(2)} = R_6^{(2)} = 1$

Tabel 3.2. Langkah Memperbaiki Policy untuk Policy $R^{(1)}$.

State	Aksi	Tes Kuantitas	
i	a	$C_i(a) - g(R^{(1)}) + \sum_{j=1}^{6} p_{ij}(a)\, v_j(R^{(1)})$	
2	0	0-0,2+(0,50)(0)+(0,20)(0)+(0,20)(0)+ (0,10)(0,8)	= -0,12
	1	1 - 0,2 + (1,0)(-0,8)	= 0
3	0	0 - 0,2 + (0,5)(0) + (0,25)(0) + (0,25)(0,8)	= 0
	1	1 - 0,2 + (1,0)(-0,8)	= 0
4	0	1 - 0,2 + (0,30)(0) + (0,70)(0,8)	= 0,36
	1	1 - 0,2 + (1,0)(-0,8)	= 0

Langkah 3.

Policy baru $R^{(2)}$ berbeda dengan policy $R^{(1)}$ yang terdahulu. Karena itu iterasi dilanjutkan.

Iterasi 2.

Langkah 1.

Untuk policy $R^{(2)}$ biaya rata-rata $g(R^{(2)})$ dan nilai relatif $v_i(R^{(2)})$ dapat dihitung sebagai solusi tunggal dari sistem persamaan linear

$v_1 = 0 - g + 0{,}75\, v_1 + 0{,}20\, v_2 + 0{,}05\, v_3$

$v_2 = 0 - g + 0{,}50 v_2 + 0{,}20\, v_3 + 0{,}20\, v_4 + 0{,}10\, v_5$

$v_k = 1 - g + v_6\ ,\ k = 3, 4$

$v_5 = 1 - g + v_6$

$v_6 = 1 - g + v_1$

$v_6 = 0$

Solusi dari sistem persamaan linear tersebut adalah

$g(R^{(2)}) = 0{,}172 \quad v_1(R^{(2)}) = -0{,}828 \quad v_2(R^{(2)}) = -0{,}178$

$v_3(R^{(2)}) = 0 \quad v_4(R^{(2)}) = 0 \quad v_5(R^{(2)}) = 0{,}828 \quad v_6(R^{(2)}) = 0$

Langkah 2.

Langkah memperbaiki policy untuk policy $R^{(2)}$ ditunjukkan dalam tabel 3.3.

Dari tabel 3.3 kita peroleh policy baru $R^{(3)}$ yang punya aksi

$R_1^{(3)} = R_2^{(3)} = 0,\ R_3^{(3)} = R_4^{(3)} = R_5^{(3)} = R_6^{(3)} = 1$

Langkah 3.

Policy baru $R^{(3)}$ sama dengan policy $R^{(2)}$, maka biaya rata-rata optimal.

Karena jika kita lanjutkan iterasinya, hasilnya persis sama dengan iterasi 2 seperti terlihat di bawah ini.

Iterasi 3.

Langkah 1.

Untuk policy $R^{(3)}$ biaya rata-rata $g\ (R^{(3)})$ dan nilai relatif $v_i\ (R^{(3)})$ dapat dihitung sebagai solusi tunggal dari sistem persamaan linear

$v_1 = 0 - g + 0{,}75\ v_1 + 0{,}20\ v_2 + 0{,}05\ v_3$

$v_2 = 0 - g + 0{,}50 v_2 + 0{,}20\ v_3 + 0{,}20\ v_4 + 0{,}10\ v_5$

$v_k = 1 - g + v_6\ ,\ k = 3, 4$

$v_5 = 1 - g + v_6$

$v_6 = 1 - g + v_1$

$v_6 = 0$

Solusi dari sistem persamaan linear tersebut adalah

$g\ (R^{(3)}) = 0{,}172 \quad v_1\ (R^{(3)}) = -\ 0{,}828 \quad v_2\ (R^{(3)}) = -0{,}178$

$v_3\ (R^{(3)}) = 0 \quad v_4\ (R^{(3)}) = 0 \quad v_5\ (R^{(3)}) = 0{,}828 \quad v_6\ (R^{(3)}) = 0$

Langkah 2.

Langkah memperbaiki policy untuk policy $R^{(3)}$ ditunjukkan dalam tabel 3.4.

Dari tabel 3.4 kita peroleh policy baru $R^{(4)}$ yang punya aksi

$R_1^{(4)} = R_2^{(4)} = 0,\ R_3^{(4)} = R_4^{(4)} = R_5^{(4)} = R_6^{(4)} = 1$

Jadi biaya rata-rata minimum dari pemeliharaan peralatan tersebut adalah 0,172.

Tabel 3.3. Langkah Memperbaiki Policy untuk Policy $R^{(2)}$.

State	Aksi	Tes Kuantitas
i	a	$C_i(a) - g(R^{(2)}) + \sum_{j=1}^{6} p_{ij}(a)\, v_j(R^{(2)})$
2	0	0 - 0,172 + (0,50)(-0,178) + (0,20)(0) + (0,20)(0) +(0,10)(0,828) = - 0,178
	1	1 - 0,172 + (1,0)(-0,828) = 0
3	0	0 - 0,172 + (0,5)(0) + (0,25)(0) + (0,25)(0,828) – 0,035
	1	1 - 0,172 + (1,0)(-0,828) = 0
4	0	1 - 0,172 + (0,30)(0) + (0,70)(0,828) = 0,408
	1	1 - 0,172 + (1,0)(-0,828) = 0

Tabel 3.4. Langkah Memperbaiki Policy untuk Policy $R^{(3)}$.

State	Aksi	Tes Kuantitas
i	a	$C_i(\{a\}) - g(R^{(3)}) + \sum_{j=1}^{6} p_{ij}(\{a\})\, v_j(R^{(3)})$
2	0	0 - 0,172 + (0,50)(-0,178) + (0,20)(0) + (0,20)(0) + (0,10)(0,828) = -0,178
	1	1 - 0,172 + (1,0)(-0,828) = 0
3	0	0 - 0,172 + (0,5)(0) + (0,25)(0) + (0,25)(0,828) = 0,035
	1	1 - 0,172 + (1,0)(-0,828) = 0
4	0	1 - 0,172 + (0,30)(0) + (0,70)(0,828) = 0,408
	1	1 - 0,172 + (1,0)(-0,828) = 0

BAB IV

KESIMPULAN DAN SARAN

A. KESIMPULAN

Dari uraian di atas diperoleh beberapa kesimpulan:

1. Metode Iterasi-Policy (Iterasi-Kebijaksanaan) merupakan suatu metode yang efektif untuk menentukan suatu kebijaksanaan yang menghasilkan biaya rata-rata harian jangka panjang optimal (minimal).
2. Metode ini merupakan suatu pilihan bagi pengambil keputusan untuk menentukan suatu kebijaksanaan.
3. Metode ini hanya cocok untuk persoalan yang mempunyai ciri-ciri yaitu keputusan yang akan diambil sekarang hanya bergantung pada keputusan yang terakhir dan tidak bergantung pada keputusan yang sebelumnya lagi.

B. SARAN-SARAN

1. Hendaknya para pengambil keputusan menggunakan metode ini untuk menentukan suatu kebijaksanaan terhadap masalah yang sesuai dengan metode ini.
2. Bagi yang tertarik untuk mendalami metode ini, hendaknya dibuatkan suatu program komputer untuk metode ini.
3. Hendaknya selalu dicari metode lain yang lebih baik dari metode ini.

DAFTAR KEPUSTAKAAN

Chung, Kai Lai. (1967). *Markov Chain,* Berlin, Springer-Verlag.

Osaki, Shunji. (1992). *Applied Stochastic System Modelling,* Berlin, Springer-Verlag.

Parzen, Emanuel. (1962). *Stochastic Processes,* San Fransisco, Holden-Day. Inc.

Ross, Seldon M. (1983). *Stochastic Processes.* New York, John Willey & Sons. Inc.

Tijms, Henk C. (1986). *Stochastic Modelling and Analysis.* New York, John Willey & Sons. Inc.